Heavy Tails and Long Range Dependence Conference in honor of Gennady Samorodnitsky's 60th birthday Paris, June 20-22 2017

Random manifolds and Gaussian processes

Robert J. Adler
Technion – Israel Institute of Technology, Israel

Although the topic has a long history, the study of *Random Manifolds and their Topology* has become a centre of major interest and activity over the past few years, mainly because of its critical importance to the burgeoning field of Applied Topology.

In this lecture, I shall present a (somewhat eclectic) collection of problems relating to random manifolds, touching lightly on the older material, and expanding on the new results and their challenges. The older material will be familiar in content and context to probabilists, but the newer material, while easy to understand at an intuitive level, requires an investment in the tools of both Algebraic and Differential Topologies for a full understanding.

The main aim of the talk will be to sufficiently excite my probabilistic friends with the intuitive (and counter-intuitive, but still easy to understand) results, that they will be convinced to try their own hand at proving even more new and exciting results about random manifolds.

Estimation of infinite–dimensional phase–type distributions

Mogens Bladt
University of Copenhagen, Denmark

If X has a phase–type distribution and N is any positive discrete random variable, then we say that the distribution of $X \cdot N$ belongs to the class of NPH distributions. Such distributions preserve the tractability and generality of phase–type distributions (often allowing for explicit solutions to stochastic models and being dense in the class of distributions on the positive reals) but with a different tail behaviour which is basically dictated by the tail of N. We thereby gain a tool for specifying distributions with a "body" shaped by X and with a tail defined by N. After reviewing the construction and basic properties

of distributions from the NPH class, we will consider the problem of their estimation. To this end we will employ the EM algorithm, using a similar method as for finite–dimensional phase–type distributions. We consider the the fitting of a NPH distribution to observed data, (left-,right and interval-) censored data, theoretical distributions, histograms, and a couple of examples.

This is a joint work with Bo Friis Nielsen and Gennady Samorodnitsky on the original part and with Leonardo Rojas-Nandayapa on the estimation.

From random matrices to long range dependence

Arijit Chakrabarty
Indian Statistical Institute, Kolkata, India

Random matrices whose entries come from a stationary Gaussian process are studied. It is shown that the limiting spectral distribution is determined by the absolutely continuous component of the spectral measure of the stationary process, a phenomenon resembling that in the situation where the entries of the matrix are i.i.d. On the other hand, the discrete component contributes to the limiting behaviour of the eigenvalues in a completely different way.

The random matrix results obtained are used to understand when a free convolution of two measures is absolutely continuous with respect to the Lebesgue measure. It is shown that if the support of a probability measure is contained in the positive half line, and is bounded away from zero, then its free multiplicative convolution with the semicircle law is absolutely continuous. For the proof, a result concerning the Hadamard product of a deterministic matrix and a scaled Wigner matrix is needed.

This talk is based on joint works with Rajat Subhra Hazra and Deepayan Sarkar.

Models for time series of counts with shape constraints

Richard A. Davis
Columbia University, USA

In recent years there has been growing interest in modelling time series of counts. Many of the formulated models for count time series are expressed via a pair of generalized state-space equations. In this set-up, the observation equation specifies the conditional distribution of the observation Y_t at time t given a *state-variable* X_t. For count time series, this conditional distribution is usually specified as coming from a known parametric family such as Poisson, negative binomial, etc. To relax this formal parametric framework, we introduce a shape constraint into the one-parameter exponential family. This essentially amounts to assuming that the *reference measure* is log-concave. In this fashion, we are

able to extend the class of observation-driven models studied in Davis and Liu (2016). Under this formulation, there exists a stationary and ergodic solution to the state-space model. In this new modeling framework, we consider the inference problem of estimating both the parameters of the mean model and the log-concave function, corresponding to the reference measure. We then compute and maximize the likelihood function over both the parameters associated with the mean function and the reference measure subject to a concavity constraint. The estimator of the mean function and the conditional distribution are shown to be consistent and perform well compared to a full parametric model specification. The finite sample behavior of the estimators are studied via simulation and two empirical examples are provided to illustrate the methodology.

This is a joint work with Jing Zhang, Columbia University.

Some problems in the domain of online social networks

Souvik Ghosh
LinkedIn, USA

LinkedIn is the world's largest professional network with more than 500M members. LinkedIn helps members connect to other members, jobs and professional opportunities via recommendations. These large scale recommendation systems are often computationally expensive. We will discuss how a member-to-member connection recommendation system works and how it affects the growth of the network. We also propose tractable generative models that mimic the social networks and discuss some interesting properties of these graphs.

On Turings formula and the estimation of the missing mass

Michael Grabchak
University of North Carolina at Charlotte, USA

The missing mass is the total probability of observing something that has not been observed before. In the context of ecological applications, it corresponds to the probability of observing a new species, while, in the context of authorship attribution studies, it corresponds to the probability that an author will use a word that he or she has not used before. Perhaps, the most famous estimator of the missing mass is Turings formula. In this talk, we give conditions for the consistency and asymptotic normality of its relative error. We then show that these conditions always hold when the distribution is regularly varying with index $\alpha \in (0; 1]$. We conclude by discussing finite sample bounds on the expectation of the missing mass and extend our results to general metric spaces.

The infinite swapping algorithm for sampling spin glasses and restricted Boltzmann machines

Henrik Hult
KTH Royal Institute of Technology, Sweden

The infinite swapping algorithm is a development of the parallel tempering or replica exchange algorithm for sampling from a Gibbs distribution. We will present this algorithm in the context of sampling spin glass models and restricted Boltzmann machines and comment on some mathematical challenges

This is a joint work with Pierre Nyquist and Carl Ringqvist, KTH.

Exact simulation of some operator scaling Gaussian random fields

Céline Lacaux
Laboratoire de Mathmatiques d'Avignon, University Avignon, France

Operator-scaling random fields, introduced in Biermé, et al.. (2007) [Operator scaling stable random fields. *Stoch. Proc. Appl.,* **117(3)**,312–332], satisfy an anisotropic self-similarity property, which extends the classical self-similarity property. Hence they generalize the fractional Brownian field, which is the most famous isotropic Gaussian self-similar random field. Up to now, to our best knowledge, such fields have only been defined through integral representations and their covariance functions are not known explicitly. Hence only approximate methods as spectral methods can be used to simulate them. In this talk we then introduce some operator Gaussian random fields with covariance defined as anisotropic deformations of the fractional Brownian field covariance and with stationary increments. This allows us to propose a fast and exact method of simulation based on the circulant embedding matrix method, following ideas of Stein 2002 [Fast and exact simulation of fractional Brownian surfaces. *Journal of Computational and Graphical Statistics*, **11(3)**,587–599] for fractional Brownian surfaces syntheses.

This is a joint work with Hermine Biermé, Poitiers University (France).

Valuation of liability cash flows

Filip Lindskog
Stockholm University, Sweden

We study the valuation of liability cash flows in a multi-period setting, where one-period capital requirements are defined in terms of a sequence of conditional risk measures.

The problem is motivated by inconsistencies in the solvency regulation for the insurance industry. We analyze properties of the liability value process resulting from basic economic assumptions and properties of the sequence of conditional risk measures. Moreover, we analyze the limiting liability value processes that appear when letting the number of periods tend to infinity and the period lengths tend to zero, and discuss requirements needed for this limiting procedure to make economic sense.

The presentation is based on current work and on Engsner, Lindholm and Lindskog (2017) [Insurance valuation: A computable multi-period cost-of-capital approach. *Insurance: Mathematics and Economics* **72**, 250-264].

Multivariate stable cumulative probabilities in polar form and related functions

John P. Nolan
American University, USA

Expressions are given for the probability of a multivariate stable random vector being in a region. These involve families of special functions that are also useful in other ways, e.g. fractional moments of stable laws, conditional expectation and the score functions for stable densities. For computational expressions, Zolotarev-type integrals are given.

Limit theorems for hyperbolic random geometric graph

Takashi Owada
Purdue University, USA

We shall consider a geometric graph model on the "hyperbolic" space, which is characterized by a negative Gaussian curvature. Among several equivalent models representing the hyperbolic space, we treat the most commonly used d-dimensional Poincare ball. One of the main characteristics of geometric graphs on the hyperbolic space is tree-like hierarchical structure. From this viewpoint, we discuss the asymptotic behavior of subtree counts. It then turns out that the spatial distribution of subtrees is crucially determined by an underlying curvature of the space. For example, if the space gets flatter and closer to the Euclidean space, subtrees are asymptotically scattered near the center of the Poincare ball. On the contrary, if the space becomes "more hyperbolic" (i.e., more negatively curved), the distribution of trees is asymptotically determined by those concentrated near the boundary of the Poincare ball.

This is a joint work with Yogeshwaran D., Indian Statistical Institute.

Financial market with no riskless (safe) asset

Svetlozar T. Rachev
Texas Tech University, USA

We study markets with no riskless (safe) asset. We derive the corresponding Black-Scholes-Merton option pricing equations for markets where there are only risky assets which have the following price dynamics: (i) continuous diffusions; (ii) jump-diffusions; (iii) diffusions with stochastic volatilities, and; (iv) geometric fractional Brownian and Rosenblatt motions. No arbitrage and market completeness conditions are derived in all four cases.

This talk is based on a joint work with Frank J. Fabozzi.

Fitting the linear preferential attachment model

Sidney Resnick
Cornell University, USA

Preferential attachment is a mechanism for modeling power-law behavior of the degree distributions in directed social networks. We consider methods for fitting a 5-parameter linear preferential model to network data under two data scenarios. In the case where full history of the network formation is given, we derive the maximum likelihood estimator of the parameters and show that they are strongly consistent and asymptotically normal. In the case where only a single-time snapshot of the network is available, we propose an estimation method which combines method of moments with an approximation to the likelihood. The resulting estimator is also strongly consistent and performs well compared to the MLE estimator. We illustrate both estimation procedures using simulated data, and explore the usage of this model in a real data example. At the end of the paper, we also present a semi-parametric method to model heavy-tailed features of the degree distributions of the network using ideas from extreme value theory.

This is a joint work with Tiandong Wang, Cornell; Richard Davis & Phyllis Wan, Columbia.

Stable random fields indexed by free groups

Parthanil Roy
Indian Statistical Institute, Bangalore, India

In this work, we investigate the extremal behaviour of left-stationary *symmetric* α-*stable* (SαS) random fields indexed by finitely generated free groups or rank strictly bigger than one. We begin by studying the rate of growth of a sequence of partial maxima

obtained by varying the indexing parameter of the field over balls of increasing size. This leads to a phase-transition that depends on the ergodic properties of the underlying nonsingular action of the free group but is different from what happens in the case of SαS random fields indexed by $\mathbb{Z}^d$. The presence of this new dichotomy is confirmed by the study of a stable random field induced by the canonical action of the free group on its Furstenberg-Poisson boundary with the measure being Patterson-Sullivan. This field is generated by a conservative action but its maxima grow as fast as an i.i.d. field contrary to what happens in the lattice case.

This talk is based on a joint work with Sourav Sarkar (presently in University of California, Berkeley), who carried out a significant portion of this work in his master's dissertation at Indian Statistical Institute.

Distributional compatibility for change of measures

Yi Shen
University of Waterloo, Canada

We study the compatibility for change of measures in this work. More precisely, for a given set of probability measures on a probability space and a corresponding set of probability distributions on the real line, we develop sufficient and necessary conditions for the existence of a random variable, such that under each measure given on the probability space, the distribution of this random variable coincides with the corresponding distribution on the real line, respectively. It is shown that the compatibility of these two sets of measures is equivalent to certain convex order condition on the Radon-Nikodym derivatives of the measures with respect to some reference measures. This result can be further extended to continuous stochastic processes.

This is a joint work with Jie Shen, Bin Wang and Ruodu Wang.

How does the generalized Rosenblatt process behave as its parameters approach the boundary of the set in which they are defined?

Murad S. Taqqu
Boston University, USA

The Rosenblatt process is one of the simplest non-Gaussian version of a self-similar process. Its marginal distributions are mixtures of chi-squares. The Rosenblatt process, however, involves a single critical exponent. The generalized Rosenblatt process is obtained by replacing that exponent by two different exponents living in the interior of a triangular region. What happens to the generalized Rosenblatt process as these critical

exponents approach the boundaries of the triangle? We show that on each of the two symmetric boundaries, the limit is non-Gaussian. On the third boundary, the limit is Brownian motion. The rates of convergence to some of these boundaries are also given. The situation is particularly delicate as one approaches the corners of the triangle, because the limit process will depend on how these corners are approached.

This is a joint work with Shuyang Bai.